

University of Groningen

Genetic and Microbial Associations to Plasma and Fecal Bile Acids in Obesity Relate to Plasma Lipids and Liver Fat Content

Chen, Lianmin; Schraa, Kiki; van den Munckhof, Inge C L; Ter Horst, Rob; Koehorst, Martijn; van Faassen, Martijn; van der Ley, Claude; Doestzada, Marwah; Zhernakova, Daria V; Kurilshikov, Alexander

Published in:
Cell reports

DOI:
[10.1016/j.celrep.2020.108212](https://doi.org/10.1016/j.celrep.2020.108212)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Chen, L., Schraa, K., van den Munckhof, I. C. L., Ter Horst, R., Koehorst, M., van Faassen, M., van der Ley, C., Doestzada, M., Zhernakova, D. V., Kurilshikov, A., Bloks, V. W., Groen, A. K., Human Functional Genomics Project, Riksen, N. P., Rutten, J. H. W., Joosten, L. A. B., Wijmenga, C., Zhernakova, A., Netea, M. G., ... Kuipers, F. (2020). Genetic and Microbial Associations to Plasma and Fecal Bile Acids in Obesity Relate to Plasma Lipids and Liver Fat Content. *Cell reports*, 33(1), [108212].
<https://doi.org/10.1016/j.celrep.2020.108212>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

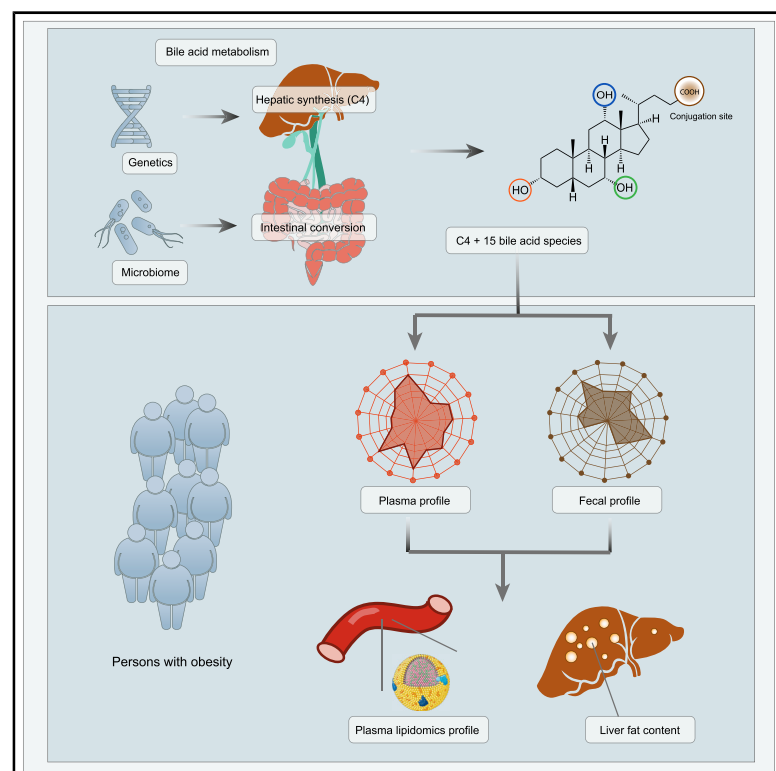
The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Genetic and Microbial Associations to Plasma and Fecal Bile Acids in Obesity Relate to Plasma Lipids and Liver Fat Content

Graphical Abstract



Authors

Lianmin Chen,
Inge C.L. van den Munckhof,
Kiki Schraa, ..., Mihai G. Netea,
Jingyuan Fu, Folkert Kuipers

Correspondence

j.fu@umcg.nl (J.F.),
f.kuipers@umcg.nl (F.K.)

In Brief

Bile acids (BAs) have been implicated in obesity-related conditions such as NAFLD and hyperlipidemia. Different human BAs exert variable biological activities. Chen et al. define genetic and microbial associations to plasma and fecal BA concentrations and composition in persons with obesity and establish their relationships with liver fat and lipid phenotypes.

Highlights

- Plasma bile acid profiles show large inter-individual variability in obesity
- Distinct genetic and microbial associations to plasma and fecal bile acid profiles
- Plasma secondary bile acids correlate with diabetes and liver fat content
- Plasma C4 correlates with features of diabetic dyslipidemia in obesity



Article

Genetic and Microbial Associations to Plasma and Fecal Bile Acids in Obesity Relate to Plasma Lipids and Liver Fat Content

Lianmin Chen,^{1,2} Inge C.L. van den Munckhof,³ Kiki Schraa,³ Rob ter Horst,³ Martijn Koehorst,^{1,4} Martijn van Faassen,⁴ Claude van der Ley,⁴ Marwah Doestzada,^{1,2} Daria V. Zhernakova,^{2,5} Alexander Kurilshikov,² Vincent W. Bloks,¹ Albert K. Groen,^{1,4,6} Human Functional Genomics Project, Niels P. Riksen,³ Joost H.W. Rutten,³ Leo A.B. Joosten,^{3,7} Cisca Wijmenga,⁸ Alexandra Zhernakova,² Mihai G. Netea,^{3,9,10,11,12} Jingyuan Fu,^{1,2,11,12,*} and Folkert Kuipers^{1,4,11,13,*}

¹Department of Pediatrics, University of Groningen, University Medical Center Groningen, Groningen 9713AV, the Netherlands

²Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen 9713AV, the Netherlands

³Department of Internal Medicine and Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Nijmegen 6500HB, the Netherlands

⁴Department of Laboratory Medicine, University of Groningen, University Medical Center Groningen, Groningen 9713AV, the Netherlands

⁵Laboratory of Genomic Diversity, Center for Computer Technologies, ITMO University, St. Petersburg 197101, Russia

⁶Department of Vascular Medicine, University of Amsterdam, Amsterdam University Medical Center, Amsterdam 1012WX, the Netherlands

⁷Department of Medical Genetics, Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca 400000, Romania

⁸University of Groningen, Groningen 9712CP, the Netherlands

⁹Department for Genomics & Immunoregulation, Life and Medical Sciences Institute, University of Bonn, Bonn 53113, Germany

¹⁰Human Genomics Laboratory, Craiova University of Medicine and Pharmacy, Craiova 200349, Romania

¹¹These authors contributed equally

¹²Senior author

¹³Lead Contact

*Correspondence: j.fu@umcg.nl (J.F.), f.kuipers@umcg.nl (F.K.)

<https://doi.org/10.1016/j.celrep.2020.108212>

SUMMARY

Bile acids (BAs) are implicated in the etiology of obesity-related conditions such as non-alcoholic fatty liver disease. Differently structured BA species display variable signaling activities via farnesoid X receptor (FXR) and Takeda G protein-coupled BA receptor 1 (TGR5). This study profiles plasma and fecal BAs and plasma 7 α -hydroxy-4-cholesten-3-one (C4) in 297 persons with obesity, identifies underlying genetic and microbial determinants, and establishes BA correlations with liver fat and plasma lipid parameters. We identify 27 genetic associations ($p < 5 \times 10^{-8}$) and 439 microbial correlations (FDR < 0.05) for 50 BA entities. Additionally, we report 111 correlations between BA and 88 lipid parameters (FDR < 0.05), mainly for C4 reflecting hepatic BA synthesis. Inter-individual variability in the plasma BA profile does not reflect hepatic BA synthetic pathways, but rather transport and metabolism within the enterohepatic circulation. Our study reveals genetic and microbial determinants of BAs in obesity and their relationship to disease-relevant lipid parameters that are important for the design of personalized therapies targeting BA-signaling pathways.

INTRODUCTION

Obesity is becoming increasingly prevalent across the globe and is associated with the development of insulin resistance, hypertension, hyperlipidemia, and non-alcoholic fatty liver disease (NAFLD), conditions collectively referred to as metabolic syndrome. Bile acids (BAs) and the BA-signaling pathways that act via the nuclear farnesoid X receptor (FXR) or the membrane-bound Takeda G protein-coupled BA receptor 1 (TGR5) have been recognized as actors in the etiologies of several components of metabolic syndrome (de Boer et al., 2018). Importantly, FXR and TGR5, as well as other receptors such as vitamin D receptor (VDR) that can be activated by specific BA-species, are expressed not only in liver and intestine, but also in adipose tissues, adrenal

glands, immune cells, and endothelial cells that are exposed to systemic BA concentrations (Chávez-Talavera et al., 2019). New therapies are emerging that target BA-signaling pathways via FXR and TGR5, particularly for NAFLD and its progressive form, non-alcoholic steatohepatitis (NASH) (de Boer et al., 2018). Furthermore, plasma concentrations of individual BAs have been proposed as biomarkers: for instance, to reflect the progression of NAFLD to NASH or the progression of pre-diabetes to type 2 diabetes (Chávez-Talavera et al., 2019; de Boer et al., 2018). These developments are, to a large extent, based on studies in pre-clinical models and on human studies in relatively small numbers of subjects (fewer than 100) (Chávez-Talavera et al., 2019).

Human BA metabolism is complex and, importantly, differs fundamentally from BA metabolism in mice and rats, the most



widely used pre-clinical models (de Boer et al., 2018). BA concentrations in the various limbs of the enterohepatic circulation (gallbladder, intestinal lumen, plasma, and stool) reflect the actions of several physiological determinants, and the human BA pool is known to display considerable inter-individual variation in both size and composition (Steiner et al., 2011). This pool consists of a mixture of primary BAs synthesized in the liver (cholic acid [CA] and chenodeoxycholic acid [CDCA]) and secondary BAs generated from CA and CDCA by intestinal microbiota (deoxycholic acid [DCA], lithocholic acid [LCA], and ursodeoxycholic acid [UDCA]), as well as a number of quantitatively minor components, such as sulfated LCA and C6-hydroxylated BAs (Molinaro et al., 2018). In steady state, hepatic synthesis of primary BAs equals fecal BA loss in order to maintain BA pool size. The contributions of the CA and CDCA synthesis pathways can thus be deduced from fecal BA composition.

Importantly, the different BA species present in the circulating BA pool show widely varying capacities to activate FXR, TGR5, and VDR and therefore to modulate metabolism and immune functions. Thus, it is important to have insight into the determinants of BA pool composition and its reflection in the systemic circulation in obese subjects if we wish to understand the contribution of endogenous BAs to obesity-related disease development, rationalize their potential use as biomarkers, and personalize future therapeutic applications of FXR and TGR5 modulators.

Since BAs are involved in the cross-talk between host genetics and gut microbiome (Fu and Kuipers, 2019), differences in genetics and gut microbiome among people with obesity will contribute to the inter-individual variations in BA concentration and composition. Yet, it is not known how genetics and gut microbiome interact in the control of BA metabolism in human obesity, how these actions translate into differences in plasma and fecal BA levels and composition, and whether these differences are actually associated with components of the metabolic syndrome in obesity.

In this study, we classified potential genetic and microbial determinants of 54 BA-related entities assessed in fasting plasma and stool samples from a relatively large cohort of 297 overweight persons with obesity-related diseases, including diabetes (12%), hypertension (60%), dyslipidemia (94%), and high liver fat content (18% with liver fat proportion >20%). This cohort thus represents the population with high metabolic syndrome risk that will likely benefit from BA-based therapies in the future. To gain insight into the potential metabolic consequences of variations in BA metabolism in these people with obesity, we also analyzed the correlations between individual BAs and metabolically defined groups of BAs and obesity-related disease phenotypes, including the levels of relevant plasma metabolites.

RESULTS

Profiles of Plasma BA Entities Show Large Variability but Modest Consistency with Stool

We assessed the concentrations of individual BA species in fasting plasma and stool samples collected from 297 and 276 participants of the 300-OB cohort, respectively. Plasma C4 levels and fecal concentrations of cholesterol, dihydro-cholesterol, and

coprostanol were also quantified. This collectively led to the definition of 58 BA-related entities (Table S1). In this study, we confined the analysis to the 54 BA entities that could be determined in >10% of the individuals (Table S1). Plasma and stool BA concentration and composition showed remarkably large inter-individual variations in this cohort (Figures 1A and 1B). Total plasma BA concentration ranged from 0.08 to 13.56 μ M (Figure 1A), a variability similar to that observed in healthy subjects (Steiner et al., 2011), while fecal BA concentrations ranged from 0.72 to 69.08 μ mol/g (dry weight). A total of 14 BA entities showed significant differences (false discovery rate [FDR] < 0.05) between males and females (e.g., the ratio of taurine-to-glycine conjugated BA in plasma), but no significant correlations were found with age (Table S1), likely related to the relatively small age range (54 to 81 years) of this cohort.

Hepatic BA synthesis in humans yields either CDCA or CA. Notably, we found that some subjects showed a strong preponderance of CDCA and its metabolites in plasma, while in other subjects, the majority of plasma BAs were derived from the CA pathway (Figure 1B). The plasma CA/CDCA ratio, also referred to as the 12 α -OH/12 α -non-OH BA ratio (Haeusler et al., 2013) and calculated as all (CA+DCA)/all (CDCA+LCA+UDCA), ranged from 0.13 to 6.82 (mean value 1.45). Surprisingly, subject-specific plasma CA/CDCA ratios were only very modestly related to the CA/CDCA ratios determined in feces ($r = 0.22$, $p = 9.4 \times 10^{-5}$; Figure 1C). The fecal CA/CDCA ratio, which reflects the ratio in which both primary BAs are produced by the liver, was 0.96 on average (range 0.15–3.22). However, and surprisingly, the fecal CA/CDCA ratio did not mimic the distinct gradient observed for plasma BA across the cohort (Figure 1B).

Correlation analyses between plasma and stool BAs revealed several distinct clusters (Figure 1C). Plasma C4 levels, which ranged from 3.92 to 835.12 μ M (Figure S1), positively correlated with total fecal BA concentrations (Figure 1D) and showed a modest positive correlation with the fecal CA/CDCA ratio ($r = 0.27$, $p = 3.3 \times 10^{-6}$; Figure 1C), suggesting a larger contribution of the CA pathway in subjects with higher total BA synthesis. In general, secondary BA entities in stool correlated positively with secondary BA entities in plasma but negatively with primary BA entities in plasma (Figure 1C). Consequently, the ratio between secondary and primary BAs in plasma and stool was associated (Figure 1E), indicating that not all primary BAs are metabolized by bacteria in the human colon. Fecal neutral sterols (i.e., cholesterol and its major bacterial metabolite coprostanol) correlated positively with total and individual BAs in feces (Figure S2). Taken together, the large inter-individual variations in BA profiles in different limbs of the enterohepatic circulation demonstrate the importance of establishing the determinants of human plasma BA profiles if we want to design effective personalized therapies based on BA-signaling pathways for treatment of obesity-related diseases.

Genetic Determinants of BA Entities

By performing quantitative trait locus (QTL) mapping for 54 BA entities, after controlling for confounding factors (age, sex, BMI, diseases, medications, and microbiome), we identified 27 independent genetics-BA associations (baQTLs) for 23 BA entities (14 plasma, 9 fecal) at a genome-wide significance level

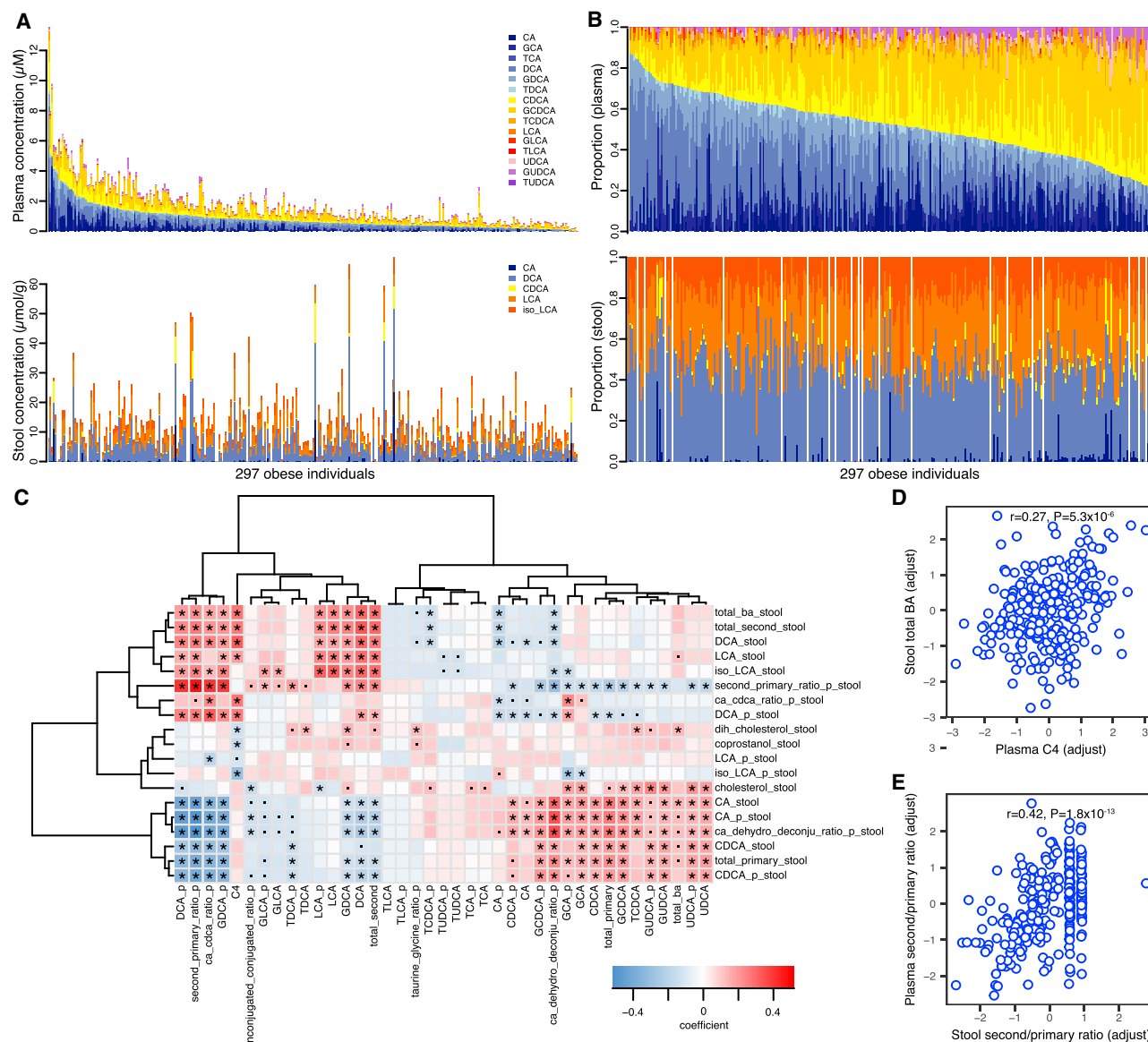


Figure 1. Profiles of Fasting Plasma and Stool BA Entities Showed Large Inter-individual Variability and Modest Consistency in the 300-OB Cohort

(A) Concentrations of fasting plasma (n = 297) and stool (n = 276) BAs in individual subjects.
 (B) Proportions of fasting plasma (n = 297) and stool (n = 276) BAs in individual subjects. Each bar represents a single individual, and the order for upper and lower panels is the same. Colors represent different BAs. White bars represent missing stool samples. UDCA and hyodeoxycholic acid were detectable in small quantities in the feces of 21% and 1.5% of the subjects, respectively, and these two BAs are not included in the figure.
 (C) Spearman correlations between plasma and stool BA entities (n = 276). Stars represent significant correlations at FDR < 0.05. Dots represent significant correlations at p < 0.05.
 (D) Spearman correlation between plasma C4 and total BAs in stool (n = 276).
 (E) Spearman correlation between secondary/primary BA ratios in plasma (n = 297) and stool (n = 276).

of $p < 5 \times 10^{-8}$ (Figure 2A; Table S2). Notably, an intron variant (rs2121703) of transporter gene Solute Carrier Family 35 Member C1 (SLC35C1), which encodes a cytidine diphosphate (CDP)-fucose transporter that regulates Wingless/Int-1 (Wnt) signaling (Deng et al., 2020), associated with plasma concentrations of DCA (Figure 2B). Another intergenic variant (rs1968543) of transporter gene Solute Carrier Family 39 Member A8

(SLC39A8, intergenic variant), which is involved in zinc and manganese transport (Park et al., 2018), associated with plasma concentrations of TDCA (Figure 2C). In addition, four zinc-finger-protein-related genetic variants (ZFN569, ZFN250, ZFN706, and JAZF1) were found to associate with five BA entities, particularly fecal BA and cholesterol content (Figure 2A; Table S2). For example, the JAZF1-related intron variant, rs10255700, was

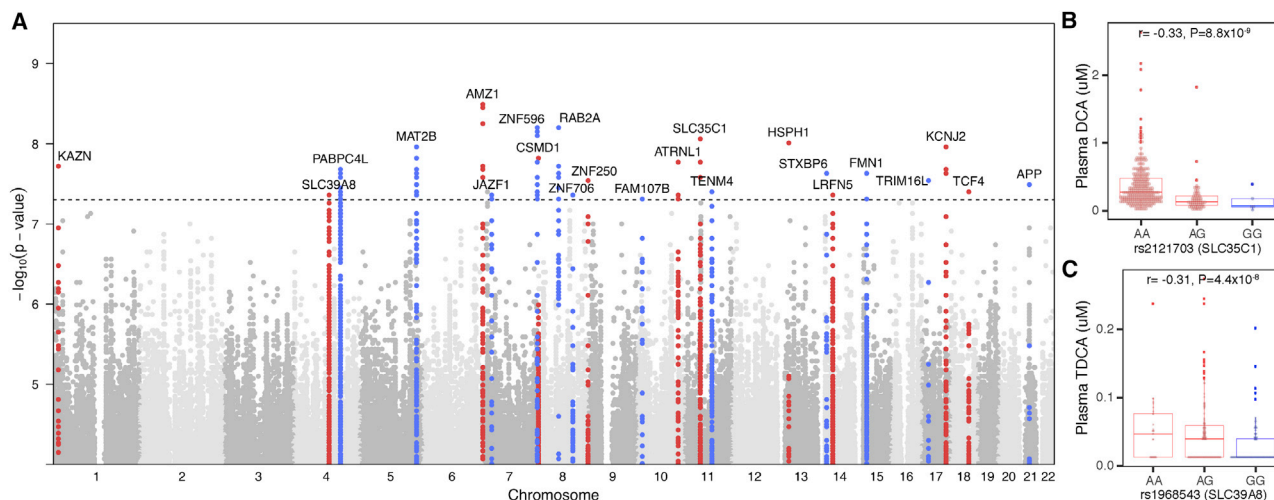


Figure 2. Genetic Determinants of BA Entities

(A) Manhattan plot showing the 27 independent ($LD\ r^2 < 0.05$, 500 kb) genetic-BA-entity associations at genome-wide significant level ($p < 5 \times 10^{-8}$). LD blocks (100 kb) are highlighted in red or blue. The representative gene of each top independent SNP is given in the plot. Names of genes and their proposed functions are given in Table S3.

(B) The SLC35C1 intron variant rs2121703 associates with plasma DCA ($n = 297$). Boxplots show medians and the first and third quartiles (the 25th and 75th percentiles) of plasma DCA concentration. The upper and lower whiskers extend the largest and smallest value no further than $1.5 \times IQR$, respectively. Outliers are plotted individually.

(C) The SLC39A8 intergenic variant rs1968543 associates with plasma TDCA ($n = 297$). Boxplots show medians and the first and third quartiles (the 25th and 75th percentiles) of plasma TDCA concentration. The upper and lower whiskers extend the largest and smallest value no further than $1.5 \times IQR$, respectively. Outliers are plotted individually.

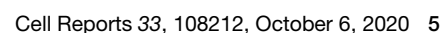
found to associate with stool cholesterol levels (Table S2). Moreover, we observed that an intergenic variant (rs55638783) of the gene encoding the human inward rectifier potassium channel Kir2.1 (KCNJ2), a lipid-gated ion channel linked to ventricular arrhythmia (Kimura et al., 2012), was associated with the concentration and proportion of plasma TDCA (Figure 2A; Table S2). Apart from these genetic determinants, we also checked genetic variants at the loci of 63 confirmed BA genes encoding enzymes, transporters, and transcription factors (Supplemental Information) and found 30 independent associations at $p < 1 \times 10^{-5}$ that related to 14 of these genes (Table S2). For instance, two independent intron variants, rs7395581 and rs10838681, related to *LXR α* (*NR1H3*), were associated with plasma DCA and total secondary BAs, respectively. Intergenic variant rs9586012, related to *SLC10A2*—which encodes the ileal apical sodium-dependent BA transporter (ASBT)—was associated with total primary BA and CDCA in stool (Table S2). Observed genetic determinants of BA entities highlight the importance of taking genetic background difference into account for developing personalized BA therapies in obesity.

Microbial Determinants of BA Entities

We next set out to identify microbial factors correlated to BA entities and found a total of 439 BA-microbial correlations for 45 BA entities at $FDR < 0.05$ (Figure 3A; Table S3). In particular, 44 BA entities correlated with 61 bacterial species (201 correlations in total; Table S3), and 30 BA entities correlated with 112 bacterial pathways (238 associations in total; Figure S3; Table S3). Both genetic associations and microbial correlations were detected

for 16 BA entities. We therefore investigated the association between BA loci and the gut microbiome (Table S4) and observed that genetics and microbiome likely play additive roles in shaping BA entities.

We found several microbial species that clustered together and showed strong positive correlations with secondary BA entities and negative correlations with primary BA entities. These included well-known microbial species able to mediate conversion of primary BAs into secondary BAs (e.g., *Eubacterium hallii* [Udayappan et al., 2016], *Ruminococcus torques* [Tonin and Arends, 2018], and *Dorea formicigenerans* [Liu et al., 2018]; Figure 3A). Additionally, many microbial correlations to BA entities became evident in this analysis. The recently identified liver-fat-content-related species *Ruminococcus sp_5_1_39BFAA* (Kurilshikov et al., 2019), for example, was positively correlated with the proportion of GDCA and the secondary/primary BA ratio in plasma (Figure 3A). We further annotated the genome of this species and identified a 3D structure of the amino acid sequence, which is encoded by a gene (g000664, from 63705 to 64702 bp) that shows high homology (global mmmmodel quality estimation, GMQE = 0.79) with cholesterylglucuronide hydrolase (Figure 3B), the microbial enzyme responsible for BA deconjugation. Moreover, we found the abundance of the anti-NAFLD species *Faecalibacterium prausnitzii* (Munukka et al., 2017) to be negatively correlated with multiple BA entities, particularly with stool levels of iso-LCA (Figures 3A and 3C). Intriguingly, the CDP-diacylglycerol biosynthesis pathway that is involved in membrane formation in this species was negatively associated with stool iso-LCA concentration (Figure 3D), suggesting that



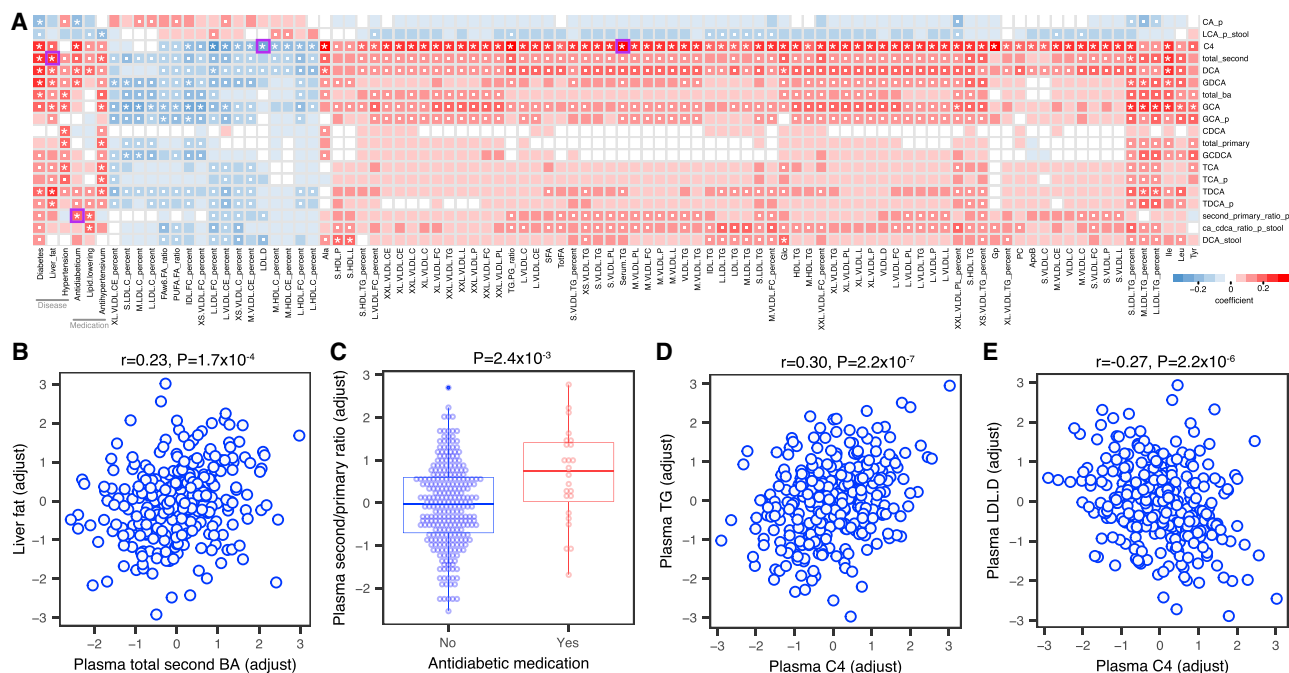


Figure 4. BA Entities Associate with Diseases, Medications, and Plasma Metabolites

- (A) The 18 disease-BA, 15 medication-BA, and 111 metabolite-BA correlations identified at FDR < 0.05. Stars represent significant correlations at FDR < 0.05. Dots represent significant correlations at $p < 0.05$.
- (B) Plasma total secondary BAs correlated with liver fat content (Spearman correlation, $n = 267$).
- (C) Difference in the plasma secondary/primary BAs ratio between antidiabetic medication users (red) and non-users (blue) (Wilcoxon test, $n = 297$).
- (D) Plasma C4 concentration correlates with plasma total triglyceride concentration (Spearman correlation, $n = 297$).
- (E) Plasma C4 concentration correlates with mean diameter of LDL particles (Spearman correlation, $n = 297$).

BA entities and plasma metabolites formed four distinct clusters (Figure 4A). The majority of the strong correlations were related to triglyceride-containing very low-density lipoproteins (VLDLs), and 48 of these showed positive correlations with plasma C4 levels (Figure 4A; Table S5), indicating that hepatic BA synthesis plays an important role in the regulation of plasma triglyceride levels in obese subjects. This is illustrated in Figure 4D, which shows the relationship between plasma C4 and total triglyceride concentrations. In addition, we observed negative relationships between plasma C4 and parameters of plasma cholesterol content (e.g., the median diameter of LDL particles, the free cholesterol-total lipid ratio in large LDL particles, and mean diameter for LDL particles; Figure 4E; Table S5). We did not observe any significant correlations between the 12α -OH/ 12α -non-OH BA ratio and plasma lipid parameters at FDR < 0.05. Taken together, however, these results indicate a close link between BA and lipid metabolism in obesity.

DISCUSSION

In this comprehensive multi-omics analysis of BA metabolism in 297 individuals, we found that the profiles of BA entities in fasting plasma and stool samples showed remarkably large inter-individual variations in the 300-OB cohort, which comprises overweight individuals aged 54 years or older (Kurilshikov et al., 2019), a targeted population with high metabolic syndrome risk

that may benefit from BA-based therapies in the future. Variability in plasma BA concentrations (Steiner et al., 2011) and C4 levels (Gälgman et al., 2011) has been reported as a feature of human BA metabolism in healthy subjects, but the underlying causes have remained unexplained. To our surprise, the large intra-individual differences across people with obesity between BAs derived from the CA and the CDCA pathway that we observed in plasma were not mirrored in stool. Yet, we found that plasma levels of C4, an established proxy of hepatic BA synthesis (Gälgman et al., 2005), were positively correlated with fecal BA concentrations. The variable contribution of CA- versus CDCA-derived BA species in feces across the cohort, ranging from 0.15 to 3.22, can therefore be interpreted to reflect the contributions of both synthesis pathways to the circulating BA pool in the individual participants. Of note, the average CA/CDCA ratio of 0.96 found in our study corresponds reasonably well to the average ratios determined by the stable isotope dilution technique in obese subjects (1.13) (Brufau et al., 2010), lean subjects on a standardized diet (1.50) (Bisschop et al., 2004), and healthy subjects on their regular diet (1.14) (Koopman et al., 1988). However, there was only a very weak relationship between the CA/CDCA ratios found in feces and those in plasma of individual participants (Figure 1B). Consequently, the large variability in plasma BA composition between subjects must be largely due to inter-individual differences in metabolic and/or transport processes that occur during enterohepatic cycling of BAs. So far,

the nature and contributions of the potential processes involved have remained unexplored, and due to the complexity of human BA metabolism, their identification requires a systems-based approach like that applied in this study.

Recently, [Kemmis et al. \(2019\)](#) performed QTL analyses on the fecal microbiome and plasma and fecal BA profiles of a diversity-outbred mouse population, a heterogeneous population derived from eight founder strains that individually harbor distinct microbial communities. Their study revealed several QTLs associated with variations in bacterial (16S sequencing) and BA profiles, with 17 loci defined as shared QTLs associating with both microbial and BA traits. However, human and murine BA metabolisms differ fundamentally in several aspects, including mouse-/rat-specific C6 hydroxylation reactions that generate muricholic acids; hepatic rehydroxylation of secondary BAs upon their return to the liver; and the murine habit of coprophagy, which causes re-entry of bacteria and (secondary) BA into the enterohepatic system ([Fu and Kuipers, 2019](#)). Our data show that in humans with obesity, both genetics and the gut microbiome relate to plasma and stool BA entities and likely play additive roles in shaping these entities.

Genetic associations to BA entities found in the current study revealed potential genetic determinants of human BA metabolism in obesity. For instance, among the 27 genetic-BA associations, we identified two independent SNPs (rs2121703 and rs1968543) related to the transporter genes *SLC35C1* (intron variant) and *SLC39A8* (intergenic variant) that were associated with plasma DCA and TDCA concentrations, respectively. These transporters are proposed to be involved in transport of GDP-fucose ([Deng et al., 2020](#)) and zinc/manganese ([Park et al., 2018](#)), respectively. How these transporters might modulate plasma levels of secondary BAs is not known and requires functional follow-up. Interestingly, [Ahmad et al. \(2020\)](#) recently reported that the non-12 α hydroxylated secondary species LCA induces the expression of the ileal manganese efflux transporter *SLC30A10* via activation of VDR, which supports the existence of a functional link between BA and manganese metabolism. Additionally, we found four zinc-finger-protein-related genetic variants (*ZFN569*, *ZFN250*, *ZFN706*, and *JAZF1*) to associate with five fecal BA entities, including fecal DCA and CDCA contents. Zinc finger proteins have been recognized to have diverse functions ([Laity et al., 2001](#)) and may potentially impact BA metabolism by regulating cholesterol or lipid metabolism. We found the *JAZF1*-related intron variant rs10255700 to associate with stool cholesterol levels ([Table S2](#)). Overexpression of *JAZF1* in apoE-deficient mice has been reported to decrease serum cholesterol levels and hepatic cholesterol synthesis by inhibiting CREB-dependent *HMGCR* promoter transcriptional activity ([Li et al., 2014](#)). We also observed that an intergenic variant of the gene encoding the human inward rectifier potassium channel Kir2.1 (*KCNJ2*), a lipid-gated ion channel linked to ventricular arrhythmia ([Kimura et al., 2012](#)), associated with plasma TDCA concentration and proportion. These secondary BAs have been reported to impact cardiac mitochondria energetics ([Ferreira et al., 2005](#)).

We also checked for genetic variants localized in or near (250 kb) 63 confirmed BA-metabolism-related genes at a lower stringency ($p < 1 \times 10^{-5}$) and confirmed BA associations with relevant transcription factors (e.g., *LXR α* , *RXR α* , *VDR*) with enzymes

involved in BA synthesis (*CYP7B1*, *AKR1D1*) and with hepatic and intestinal transporters (*ABCC1* and *SLC10A2/ASBT*). Interestingly, the rs9586012 intergenic variant related to the ileal BA transporter ASBT was associated with the fecal content of primary BAs, suggesting increased influx of primary BAs into the colon and/or decreased microbial BA deconjugation/dehydroxylation activity. In their mouse study, [Kemmis et al. \(2019\)](#) identified a QTL near *Slc10a2/Asbt* that was associated with both the abundance of *Turicibacter* sp., which shows BA-deconjugating activity, and the plasma concentration of unconjugated CA.

Microbial correlations to BA entities that we found not only confirm previous microbial BA determinants ([Liu et al., 2018](#); [Tonin and Arends, 2018](#); [Udayappan et al., 2016](#)), but also include extra correlations, such as those with *Ruminococcus* sp_5_1_39BFAA and *Faecalibacterium prausnitzii*, that may reflect conversion of BAs by the bacterium or, conversely, modulation of the bacterium by the actions of (bacteriostatic) BAs. It is important to note that we only focused on the most prominent BA species. In the future, it would be interesting to also assess the determinants of (much) less abundant secondary BA species, such as the LCA metabolites 3-oxo-LCA and isoallo-LCA that were recently shown to control T cell differentiation in the colonic lamina propria and, hence, modulate adaptive immunity ([Hang et al., 2019](#)).

We also examined to what extent BA entities correlated with obesity-related diseases, including liver fat content, diabetes, hypertension, and dyslipidemia (after controlling for age and sex). This analysis identified seven BA correlations with diabetes, six BA associations with liver fat content, and five BA correlations with hypertension at FDR < 0.05. The majority of these actually related to bacteria-derived BA entities, particularly for total secondary BA, DCA, GDCA, and TDCA. BA-liver-fat-content correlations support a role for BAs in the etiology of NAFLD in these obese subjects. Preclinical studies have shown that BAs can impact the development of NAFLD by modulating lipogenesis via activation of FXR ([de Boer et al., 2018](#)) and the SREBP1c-pathway ([Watanabe et al., 2004](#)), as well as by suppressing VLDL production ([Lin et al., 1996](#)), while elevated levels of DCA have previously been related to NAFLD and NASH in humans ([Jiao et al., 2018](#)). Intriguingly, we observed that *Ruminococcus* sp_5_1_39BFAA—a species expressing the gene encoding cholesterylglucuronide hydrolase ([Figure 3B](#)), which catalyzes BA deconjugation—correlated not only with bacterial BA entities, but also with hepatic fat content ([Kurilshikov et al., 2019](#)). These results support the hypothesis that *Ruminococcus* sp_5_1_39BFAA might contribute to NAFLD by co-regulating intestinal BA metabolism. Additionally, we observed the abundance of *F. prausnitzii*, a butyrate producer shown to reduce hepatic fat content in mice ([Munukka et al., 2017](#)), to be negatively associated with stool iso-LCA concentration and liver fat content. The CDP-diacylglycerol biosynthesis pathway of this species, which plays an important role in the formation of bacterial membranes, was negatively associated with stool iso-LCA concentration, suggesting that iso-LCA may act as antimicrobial agent that inhibits the growth of these anti-NAFLD bacteria.

Importantly, our study identified 111 links between BA entities and disease-relevant plasma lipoprotein characteristics, which supports an important role for BAs in the regulation of human

lipid metabolism. The positive correlation between plasma C4 and triglyceride levels is in agreement with the results of Gälman et al. (2011) in a cohort of 435 healthy volunteers with an age range from 20 to 89 years. In addition, Brufau et al. (2010) showed that induction of hepatic BA synthesis, as quantified by stable isotope dilution, was linearly related to plasma triglyceride levels in obese subjects treated with the BA sequestrant colessevelam. Conversely, certain forms of familial hypertriglyceridemia were found to be characterized by high BA synthesis, possibly due to inborn defects of intestinal BA reabsorption (Angelin et al., 1987), and treatment with CDCA, which suppresses endogenous BA synthesis, has been shown to reduce plasma triglycerides in hypertriglyceridemic subjects (Angelin et al., 1987). Interestingly, we found fasting C4 to be positively correlated with differently sized triglyceride-containing VLDLs, suggesting that the relationship originates in the VLDL production process rather than in the modulation of the VLDL de-lipidation cascade in plasma. Increased production of VLDL by the liver has been observed in obese subjects, likely related to hepatic insulin resistance (Pramfalk et al., 2016), but BAs themselves have also been shown to suppress the hepatic VLDL production process (Lin et al., 1996). Interference with FXR-mediated modulation of *de novo* lipogenesis in the liver may contribute to these effects (Sjöberg et al., 2017).

In conclusion, this systematic approach has revealed the contributions of genetics and the microbiome to plasma and stool BA entities in human obesity. The surprisingly large inter-individual variations observed in plasma BA entities highlight that both genetic and microbial determinants should be taken into account in designing studies to evaluate future treatment strategies that target BA-signaling pathways. In particular, the broad correlations of microbiome-derived secondary BA entities with hepatic fat content, diabetes, and hypertension observed in these obese subjects, in conjunction with very recent reports that link secondary BAs to intestinal immunity (Hang et al., 2019), underline the importance of our microbial “second genome” in determining the biological activity of an individual’s BA pool. In addition, the strong correlation of plasma C4 with circulating triglyceride-containing lipoproteins and specific LDL features in obese subjects warrants dedicated evaluation when interfering with human BA metabolism in the context of NAFLD therapies. Indeed, subjects with high initial BA synthesis rates (high C4) may respond more pronouncedly to pharmacological FXR agonists with respect to plasma and hepatic lipid parameters than subjects with low synthesis rates. In view of the very large inter-individual variations in BA concentration and composition present in obese subjects (i.e., a target population for pharmacological interventions with FXR or TGR5 modulators), it will be important to define measures to assess basal activation status of FXR and TGR5 in humans to allow for the design of effective personalized therapies.

We acknowledge several limitations in present study. First, the 300-OB cohort comprises participants of only Dutch ethnicity. The reported results are thus likely biased toward a region-specific genetic and microbial background. Second, this study represents an association/correlation analysis based on a cross-sectional study design, which means that the underlying causalities and mechanisms of action remain unexplored. Longi-

tudinal studies and, particularly, functional studies are thus essential to reveal the underlying mechanisms of the reported associations. Third, although we have taken several confounding factors into account, additional internal and external confounding factors (e.g., physical status, mental health, and environmental exposures) may have influenced outcomes. Fourth, given the fact that it is practically impossible to collect and mix all feces produced during 3 consecutive days, as needed for accurate determination of BA synthesis from fecal output, fecal BA analysis relied on a sample taken from one single feces collection, which might thus not exactly represent the ratio in which primary BAs are being produced by the liver.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead Contact
 - Materials Availability
 - Data and Code Availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
 - Human participants
- **METHOD DETAILS**
 - Bile acids and their related entities
 - Liver fat content quantification
 - Genotyping
 - Microbial species and pathway abundances
 - Genomic annotation of *Ruminococcus* sp_5_1_39BFAA
 - Plasma metabolite measurement
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Genome-wide associations
 - Microbiome-wide correlation
 - Independent effects of genetics and microbiome
 - Correlation between disease and medication phenotypes
 - Correlation with disease-relevant plasma metabolites

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.celrep.2020.108212>.

ACKNOWLEDGMENTS

We thank all the volunteers in the 300-OB cohort of the Human Functional Genomics Project (HFGP) for their participation and the project staff for their help and management. We thank Kate McIntyre for editing the text. This project was supported by the Netherlands Heart Foundation (IN-CONTROL CVON grant 2012-03 to M.G.N., L.A.B.J., F.K., A.Z., and J.F. and IN-CONTROL CVON grant 2018-27 to M.G.N., N.P.R., L.A.B.J., F.K., A.Z., and J.F.); the Noaber Foundation (F.K.); the Netherlands Organization for Scientific Research (NWO) (NWO-VIDI 864.13.013 to J.F.; NWO-VIDI 016.178.056 to A.Z.; NWO Spinoza Prize SPI 94-212 to M.G.N.; NWO Spinoza Prize SPI 92-266 to C.W.; NWO-VENI 194.006 to D.V.Z.; the NWO Gravitation Netherlands Organ-on-Chip Initiative to J.F. and C.W.; and the NWO Gravitation Exposure-NL 024.004.017 to J.F., A.K., and A.Z.); the European Research Council (ERC) (FP7/2007-2013/ERC Advanced Grant Agreement 2012-322698 to C.W., ERC Consolidator Grant 310372 to M.G.N., and ERC Starting Grant

715772 to A.Z.); the Stiftelsen Kristian Gerhard Jebsen Foundation (Norway) to C.W.; the RuG Investment Agenda Grant Personalized Health to C.W.; and the Foundation De Cock-Hadders grant (20:20-13) to L.C. A.Z. holds a Rosalind Franklin Fellowship from the University of Groningen. L.C. is supported by a joint fellowship from the University Medical Center Groningen and China Scholarship Council (CSC201708320268). L.A.B.J. is supported by a Competitiveness Operational Programme grant of the Romanian Ministry of European Funds (HINT, P_37_762, MySMIS 103587). This research also received funding under the European Union Seventh Framework Program. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

J.F., M.G.N., and F.K. conceptualized and managed the study. L.C., I.C.L.v.d.M., K.S., R.t.H., M.K., M.v.F., M.D., D.V.Z., A.K., C.v.d.L., A.K.G., N.P.R., J.H.W.R., and L.A.B.J. collected the samples and generated the data. L.C. analyzed the data. L.C., J.F., and F.K. drafted the manuscript. L.C., I.C.L.v.d.M., K.S., R.t.H., M.K., V.W.B., M.v.F., M.D., C.v.d.L., D.V.Z., A.K., A.K.G., N.P.R., J.H.W.R., L.A.B.J., C.W., A.Z., M.G.N., J.F., and F.K. reviewed and edited the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 25, 2020

Revised: July 31, 2020

Accepted: September 9, 2020

Published: October 6, 2020

REFERENCES

Ahmad, T.R., Higuchi, S., Bertaggia, E., Hung, A., Shanmugarajah, N., Guilz, N.C., Gamarra, J.R., and Haeusler, R.A. (2020). Bile acid composition regulates the manganese transporter Slc30a10 in intestine. *J. Biol. Chem.* 295, 12545–12558.

Angelin, B., Hershon, K.S., and Brunzell, J.D. (1987). Bile acid metabolism in hereditary forms of hypertriglyceridemia: evidence for an increased synthesis rate in monogenic familial hypertriglyceridemia. *Proc. Natl. Acad. Sci. USA* 84, 5434–5438.

Bisschop, P.H., Bandsma, R.H., Stellaard, F., ter Harmsel, A., Meijer, A.J., Sauerwein, H.P., Kuipers, F., and Romijn, J.A. (2004). Low-fat, high-carbohydrate and high-fat, low-carbohydrate diets decrease primary bile acid synthesis in humans. *Am. J. Clin. Nutr.* 79, 570–576.

Brufau, G., Stellaard, F., Prado, K., Bloks, V.W., Jonkers, E., Boverhof, R., Kuipers, F., and Murphy, E.J. (2010). Improved glycemic control with colestevam treatment in patients with type 2 diabetes is not directly associated with changes in bile acid metabolism. *Hepatology* 52, 1455–1464.

Chávez-Talavera, O., Haas, J., Grzych, G., Tailleux, A., and Staels, B. (2019). Bile acid alterations in nonalcoholic fatty liver disease, obesity, insulin resistance and type 2 diabetes: what do the human studies tell? *Curr. Opin. Lipidol.* 30, 244–254.

Chen, L., Garmaeva, S., Zhernakova, A., Fu, J., and Wijmenga, C. (2018). A system biology perspective on environment-host-microbe interactions. *Hum. Mol. Genet.* 27, R187–R194.

Chen, L., Collij, V., Jaeger, M., van den Munckhof, I.C.L., Vich Vila, A., Kurilshikov, A., Gacesa, R., Sinha, T., Oosting, M., Joosten, L.A.B., et al. (2020). Gut microbial co-abundance networks show specificity in inflammatory bowel disease and obesity. *Nat. Commun.* 11, 4018.

de Boer, J.F., Bloks, V.W., Verkade, E., Heiner-Fokkema, M.R., and Kuipers, F. (2018). New insights in the multiple roles of bile acids and their signaling pathways in metabolic control. *Curr. Opin. Lipidol.* 29, 194–202.

Deng, M., Chen, Z., Tan, J., and Liu, H. (2020). Down-regulation of SLC35C1 induces colon cancer through over-activating Wnt pathway. *J. Cell. Mol. Med.* 24, 3079–3090.

Eggink, H.M., Oosterman, J.E., de Goede, P., de Vries, E.M., Foppen, E., Koe-horst, M., Groen, A.K., Boelen, A., Romijn, J.A., la Fleur, S.E., et al. (2017). Complex interaction between circadian rhythm and diet on bile acid homeostasis in male rats. *Chronobiol. Int.* 34, 1339–1353.

Ferreira, M., Coxito, P.M., Sardão, V.A., Palmeira, C.M., and Oliveira, P.J. (2005). Bile acids are toxic for isolated cardiac mitochondria: a possible cause for hepatic-derived cardiomyopathies? *Cardiovasc. Toxicol.* 5, 63–73.

Franzosa, E.A., McIver, L.J., Rahnava, G., Thompson, L.R., Schirmer, M., Weingart, G., Lipson, K.S., Knight, R., Caporaso, J.G., Segata, N., and Huttenhower, C. (2018). Species-level functional profiling of metagenomes and metatranscriptomes. *Nat. Methods* 15, 962–968.

Fu, J., and Kuipers, F. (2019). Systems genetics approach reveals cross-talk between bile acids and intestinal microbes. *PLoS Genet.* 15, e1008307.

Gälman, C., Angelin, B., and Rudling, M. (2005). Bile acid synthesis in humans has a rapid diurnal variation that is asynchronous with cholesterol synthesis. *Gastroenterology* 129, 1445–1453.

Gälman, C., Angelin, B., and Rudling, M. (2011). Pronounced variation in bile acid synthesis in humans is related to gender, hypertriglyceridaemia and circulating levels of fibroblast growth factor 19. *J. Intern. Med.* 270, 580–588.

Haeusler, R.A., Astiarraga, B., Camastra, S., Accili, D., and Ferrannini, E. (2013). Human insulin resistance is associated with increased plasma levels of 12 α -hydroxylated bile acids. *Diabetes* 62, 4184–4191.

Hang, S., Paik, D., Yao, L., Kim, E., Trinath, J., Lu, J., Ha, S., Nelson, B.N., Kelly, S.P., Wu, L., et al. (2019). Bile acid metabolites control T_H17 and T_{reg} cell differentiation. *Nature* 576, 143–148.

Hoogerland, J.A., Lei, Y., Wolters, J.C., de Boer, J.F., Bos, T., Bleeker, A., Mulder, N.L., van Dijk, T.H., Kuivenhoven, J.A., Rajas, F., et al. (2019). Glucose-6-Phosphate Regulates Hepatic Bile Acid Synthesis in Mice. *Hepatology* 70, 2171–2184.

Jakulj, L., van Dijk, T.H., de Boer, J.F., Kootte, R.S., Schonewille, M., Paalvast, Y., Boer, T., Bloks, V.W., Boverhof, R., Nieuwdorp, M., et al. (2016). Transintestinal Cholesterol Transport Is Active in Mice and Humans and Controls Ezetimibe-Induced Fecal Neutral Sterol Excretion. *Cell Metab.* 24, 783–794.

Jiao, N., Baker, S.S., Chapa-Rodriguez, A., Liu, W., Nugent, C.A., Tsompana, M., Mastrandrea, L., Buck, M.J., Baker, R.D., Genco, R.J., et al. (2018). Suppressed hepatic bile acid signalling despite elevated production of primary and secondary bile acids in NAFLD. *Gut* 67, 1881–1891.

Kemis, J.H., Linke, V., Barrett, K.L., Boehm, F.J., Traeger, L.L., Keller, M.P., Rabaglia, M.E., Schueler, K.L., Stapleton, D.S., Gatti, D.M., et al. (2019). Genetic determinants of gut microbiota composition and bile acid profiles in mice. *PLoS Genet.* 15, e1008073.

Kimura, H., Zhou, J., Kawamura, M., Itoh, H., Mizusawa, Y., Ding, W.G., Wu, J., Ohno, S., Makiyama, T., Miyamoto, A., et al. (2012). Phenotype variability in patients carrying KCNJ2 mutations. *Circ. Cardiovasc. Genet.* 5, 344–353.

Koopman, B.J., Kuipers, F., Bijleveld, C.M., van der Molen, J.C., Nagel, G.T., Vonk, R.J., and Wolthers, B.G. (1988). Determination of cholic acid and chenodeoxycholic acid pool sizes and fractional turnover rates by means of stable isotope dilution technique, making use of deuterated cholic acid and chenodeoxycholic acid. *Clin. Chim. Acta* 175, 143–155.

Kurilshikov, A., van den Munckhof, I.C., Chen, L., Bonder, M.J., Schraa, K., Rutten, J., Riksen, N.P., de Graaf, J., Oosting, M., Sanna, S., et al. (2019). Gut Microbial Associations to Plasma Metabolites Linked to Cardiovascular Phenotypes and Risk: A Cross-Sectional Study. *Circ. Res.* 124, 1808–1820.

Laity, J.H., Lee, B.M., and Wright, P.E. (2001). Zinc finger proteins: new insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* 11, 39–46.

Li, X., Yang, M., Wang, H., Jia, Y., Yan, P., Boden, G., Yang, G., and Li, L. (2014). Overexpression of JAZF1 protected ApoE-deficient mice from atherosclerosis by inhibiting hepatic cholesterol synthesis via CREB-dependent mechanisms. *Int. J. Cardiol.* 177, 100–110.

- Lin, Y., Havinga, R., Verkade, H.J., Moshage, H., Slooff, M.J.H., Vonk, R.J., and Kuipers, F. (1996). Bile acids suppress the secretion of very-low-density lipoprotein by human hepatocytes in primary culture. *Hepatology* 23, 218–228.
- Liu, H., Hu, C., Zhang, X., and Jia, W. (2018). Role of gut microbiota, bile acids and their cross-talk in the effects of bariatric surgery on obesity and type 2 diabetes. *J. Diabetes Investig.* 9, 13–20.
- Loh, P.R., Danecek, P., Palamara, P.F., Fuchsberger, C., A Reshef, Y., K Finucane, H., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G.R., et al. (2016). Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* 48, 1443–1448.
- Molinaro, A., Wahlström, A., and Marschall, H.U. (2018). Role of Bile Acids in Metabolic Control. *Trends Endocrinol. Metab.* 29, 31–41.
- Munukka, E., Rintala, A., Toivonen, R., Nylund, M., Yang, B., Takanen, A., Hänninen, A., Vuopio, J., Huovinen, P., Jalkanen, S., and Pekkala, S. (2017). *Faecalibacterium prausnitzii* treatment improves hepatic health and reduces adipose tissue inflammation in high-fat fed mice. *ISME J.* 11, 1667–1679.
- Park, J.H., Hogrebe, M., Fobker, M., Brackmann, R., Fiedler, B., Reunert, J., Rust, S., Tsiakas, K., Santer, R., Grüneberg, M., and Marquardt, T. (2018). SLC39A8 deficiency: biochemical correction and major clinical improvement by manganese therapy. *Genet. Med.* 20, 259–268.
- Pramfalk, C., Pavlides, M., Banerjee, R., McNeil, C.A., Neubauer, S., Karpe, F., and Hodson, L. (2016). Fasting Plasma Insulin Concentrations Are Associated With Changes in Hepatic Fatty Acid Synthesis and Partitioning Prior to Changes in Liver Fat Content in Healthy Adults. *Diabetes* 65, 1858–1867.
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069.
- Sjöberg, B.G., Straniero, S., Angelin, B., and Rudling, M. (2017). Cholestyramine treatment of healthy humans rapidly induces transient hypertriglyceridemia when treatment is initiated. *Am. J. Physiol. Endocrinol. Metab.* 313, E167–E174.
- Steiner, C., Othman, A., Saely, C.H., Rein, P., Drexel, H., von Eckardstein, A., and Rentsch, K.M. (2011). Bile acid metabolites in serum: intraindividual variation and associations with coronary heart disease, metabolic syndrome and diabetes mellitus. *PLoS ONE* 6, e25006.
- Tonin, F., and Arends, I.W.C.E. (2018). Latest development in the synthesis of ursodeoxycholic acid (UDCA): a critical review. *Beilstein J. Org. Chem.* 14, 470–483.
- Truong, D.T., Franzosa, E.A., Tickle, T.L., Scholz, M., Weingart, G., Pasolli, E., Tett, A., Huttenhower, C., and Segata, N. (2015). MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* 12, 902–903.
- Udayappan, S., Manneras-Holm, L., Chaplin-Scott, A., Belzer, C., Herrema, H., Dallinga-Thie, G.M., Duncan, S.H., Stroes, E.S.G., Groen, A.K., Flint, H.J., et al. (2016). Oral treatment with *Eubacterium hallii* improves insulin sensitivity in *db/db* mice. *NPJ Biofilms Microbiomes* 2, 16009.
- Watanabe, M., Houten, S.M., Wang, L., Moschetta, A., Mangelsdorf, D.J., Heyman, R.A., Moore, D.D., and Auwerx, J. (2004). Bile acids lower triglyceride levels via a pathway involving FXR, SHP, and SREBP-1c. *J. Clin. Invest.* 113, 1408–1418.
- Watanabe, K., Taskesen, E., van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* 8, 1826.
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303.

STAR★METHODS

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|--|---|---|
| Biological Samples | | |
| Fecal samples | This study | N/A |
| Blood samples | This study | N/A |
| Critical Commercial Assays | | |
| AllPrep DNA/RNA Mini Kit | QIAGEN | 80204 |
| Quant-iT PicoGreen dsDNA Assay | Life Technologies | P7589 |
| Nextera XT DNA Library Preparation kit | Illumina | FC-131-1096 |
| HumanCoreExome-24 BeadChip | Illumina | https://www.illumina.com |
| Infinium OmniExpress-24 BeadChip | Illumina | https://www.illumina.com |
| Nightingale Lipidomics | Nightingale | https://nightingalehealth.com |
| Bile Acid Assays | (Hoogerland et al., 2019) | PMID: 31102537 |
| Software and Algorithms | | |
| R (3.6.0) | R Foundation | http://www.r-project.org/ |
| Java (1.8.0_74) | Java | https://www.java.com/en/ |
| Michigan Imputation Server | (Loh et al., 2016) | PMID: 27694958 |
| MetaPhlan (2.7.2) | (Truong et al., 2015) | PMID: 26418763 |
| HUMAnN2 | (Franzosa et al., 2018) | PMID: 30377376 |
| Prokka (1.11) | (Seemann, 2014) | PMID: 24642063 |
| jMRUI (3.0) | http://www.jmrui.eu | http://www.jmrui.eu |
| Deposited Data | | |
| Raw metagenomics | EGA | https://www.ebi.ac.uk |
| Genotype | HFGP | http://www.humanfunctionalgenomics.org |
| Bile acid and lipid traits | HFGP | http://www.humanfunctionalgenomics.org |
| Clinical phenotypes | HFGP | http://www.humanfunctionalgenomics.org |

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to Lead Contact, Folkert Kuipers (f.kuipers@umcg.nl).

Materials Availability

The study did not generate any new reagents or materials.

Data and Code Availability

The accession number for the metagenomic sequencing data reported in this paper is deposited in the European Genome-Phenome Archive (EGA, <https://www.ebi.ac.uk/ega/home>): EGAD00001004194. Other datasets including genotype, bile acid profiles, NMR-based metabolome, and phenotypic data are available via: <http://www.humanfunctionalgenomics.org>

Analysis codes are available via:

https://github.com/GRONINGEN-MICROBIOME-CENTRE/Groningen-Microbiome/tree/master/Projects/300OB_BileAcids

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human participants

302 individuals aged 54 to 81 years were enrolled in the 300-OB cohort at the Radboud University Medical Center (Radboud UMC), Nijmegen, the Netherlands. All participants had a body mass index (BMI) ≥ 27 kg/m² at screening (mean = 30.73, median = 29.89) and the study has been approved by Radboud UMC with number NL46846.091.13 (Chen et al., 2020). Participants with a recent cardiovascular event (myocardial infarction, transient ischemic attack, stroke < 6 months), a history of bariatric surgery or bowel

resection, inflammatory bowel disease, renal dysfunction, increased bleeding tendency, use of oral or subcutaneous anti-coagulant therapy, use of thrombocyte aggregation inhibitors other than acetylsalicylic acid and carbasalate calcium, or a contra-indication for magnetic resonance imaging were excluded from the study. All participating women were postmenopausal and did not use hormonal replacement therapy. For all participants, blood samples were collected for BA, C4 and nuclear magnetic resonance (NMR)-based lipidomics measurements in the morning following an overnight fast (Kurilshikov et al., 2019). Blood samples were also used for genotyping. Fecal samples for metagenomics sequencing and BA measurements were collected the day before blood collection. The present study includes the 297 participants for whom genotype, stool microbiome, medical records and fasting plasma metabolites were available. Fasting plasma BAs were determined in 166 men and 131 women with an average age of 67.0 years (sd = 5.4). We quantified stool BA and neutral sterol concentrations in 276 of these 297 participants.

METHOD DETAILS

Bile acids and their related entities

Levels of 15 BAs and C4 concentrations in fasting plasma were quantified by LC-MS procedures, as previously described (Eggink et al., 2017; Hoogerland et al., 2019). In stool samples, we measured concentrations of 7 BAs and the neutral sterols cholesterol, dihydrocholesterol and coprostanol (Jakulj et al., 2016). We grouped BAs into the concentrations of total BAs, conjugated and unconjugated BAs, and primary and secondary BAs based on their respective biological functionalities. We also calculated the relative proportion of each BA and constructed five BA ratios to reflect hepatic and bacterial enzymatic activities: (1) the ratio of secondary versus primary BAs for microbial activity, (2) the ratio of CA species (CA, GCA, TCA, DCA, GDCA, TDCA) versus CDCA species (CDCA, GCDCA, TCDCA, LCA, GLCA, TLCA) for the relative activities of the two major hepatic BA biosynthesis pathways, (3) the ratio of unconjugated versus conjugated BAs, and (4) the ratio of taurine-conjugated versus glycine-conjugated BAs, (5) ratio between dehydrogenase cholic acid species and deconjugated cholic acid species. We also calculated the proportions of each individual BA species and of groups of BAs. This led to a total of 58 BA entities, with 54 of these present in more than 10% of the participants (Table S1). For BA concentrations under the detection limit, we replaced the missing value with half of the minimum value of that specific BA.

Liver fat content quantification

Liver fat content was quantified using localized proton magnetic resonance spectroscopy (1H-MRS). A single cubical voxel of 27 mL was positioned in the right lobe of the liver. The voxel was placed outside the biliary tree and blood vessels to avoid confounding the region of interest. A STEAM localization sequence without water suppression was used for data acquisition. To minimize relaxation effects on signal intensity, a long repetition time (TR = 3 s) and a short echo time (TE = 20 ms) were used. All MR spectra were post-processed using the jMRUI software v3.0 package with the AMARES algorithm to determine water (4.7 ppm) and methylene (1.3 ppm) resonance areas. Intrahepatic triglyceride content was expressed as a ratio of methylene signal area to the sum of the water and methylene signal areas (%).

Genotyping

Around half of the samples (n = 134) were genotyped using the Illumina HumanCoreExome-24 BeadChip Kit. All other samples (n = 168) were genotyped using the Illumina Infinium Omni-express chip. The two sets were merged by only keeping single nucleotide polymorphisms (SNPs) present in both datasets. Imputation was done using the Michigan Imputation Server with the Haplotype Reference Consortium reference panel hrc.r1.1.201691, and phasing was done by using Eagle (Loh et al., 2016). We excluded SNPs that had imputation quality $r^2 < 0.3$, failed the Hardy-Weinberg equilibrium test ($p < 1 \times 10^{-4}$), had a call rate < 99%, or had a minor allele frequency (MAF) < 10%. In this way, we obtained genotype data for 4.3 million SNPs for all participants. After imputation, principle component analysis was used to verify that there were no differences between the two datasets.

Microbial species and pathway abundances

All participants were asked to collect a fecal sample at home and to place it in their home freezer within 15 minutes after production. Participants then brought the sample in frozen state to the hospital, where the samples were placed on dry ice and transferred to the laboratory. Aliquots were then made and stored at -80°C until further processing. Fecal DNA isolation was performed using the AllPrep DNA/RNA Mini Kit (QIAGEN; cat. 80204). After DNA extraction, fecal DNA was sent to the Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA, where library preparation and whole genome shotgun sequencing were performed on the Illumina HiSeq platform. Low-quality reads were discarded from the raw metagenomic sequencing data by the sequencing facility, and reads belonging to the human genome were removed by mapping the data to the human reference genome (version NCBI37) with Bowtie2 (v.2.1.0).

The relative abundance of gut microbial taxonomic units was determined using MetaPhlan v.2.7.2 (Truong et al., 2015). For these analyses, we only included taxonomy data at species-level. Classified species present in > 10% of the samples were included for further analyses. This yielded a list of 173 species. The relative abundances of metabolic pathways were determined using the HUMAnN2 pipeline (Franzosa et al., 2018), which maps DNA/RNA reads to a customized database of functionally annotated pan-genomes. HUMAnN2 reported the abundances of gene families from the UniProt Reference Clusters (UniRef90), which were further mapped to microbial pathways from the MetaCyc metabolic pathway database. In total, 352 microbial pathways present in > 10% of samples were selected for subsequent analyses.

Genomic annotation of *Ruminococcus* sp_5_1_39BFAA

The reference genome of *Ruminococcus* sp_5_1_39BFAA was downloaded from the National Center for Biotechnology Information (NCBI) with accession number GCA_000159975.2. Prokka software (Seemann, 2014) was used to annotate the reference genome of *Ruminococcus* sp_5_1_39BFAA, which resulted in a list of genes with functional annotations. A gene (g000664) potentially encoding choloylglycine hydrolase, a BA-deconjugation enzyme, has been annotated in the genome of this species. To validate whether the classified gene has BA deconjugation abilities, we used SWISS-MODEL (Waterhouse et al., 2018) to predict its 3D structure based on the translated amino acids sequence of this gene and compared this with existing choloylglycine hydrolase structures.

Plasma metabolite measurement

A wide range of plasma metabolites were measured using NMR and Nightingale's Biomarker Analysis Platform. This platform provides measures of 231 plasma metabolome traits. Most of these are related to lipids, including total lipid concentrations, relative compositions of 14 lipoprotein subclasses, lipoprotein particle sizes, and concentrations of apolipoproteins, cholesterol, triglycerides and phospholipids. The measurements also include several glycolysis components, fatty acids, inflammation markers, ketone bodies and amino acids (Kurilshikov et al., 2019). To validate platform precision, we compared several traits with corresponding routine lipid measurements and observed a high degree of consistency (Kurilshikov et al., 2019). Finally, 225 metabolites present in more than 10% of the participants were included in subsequent analyses.

QUANTIFICATION AND STATISTICAL ANALYSIS

Genome-wide associations

To identify potential genetic determinants of BAs, we performed quantitative trait locus (QTL) mapping by calculating Spearman correlations between BA entities and SNP dosages. We also included the ratios of 7 BA measurements that were present in both plasma and stool (the ratios of CA, LCA, DCA, CDCA, total BA, total primary BA and total secondary BA between plasma and stool samples) in QTL mapping to identify potential BA receptors. Since a number of confounding factors can impact BAs, we first inverse-rank-transformed BA entities and then stepwise adjusted for these confounding factors (Table S1) – age, sex, BMI, disease (diabetes, hypertension and dyslipidemia), medication use (antidiabetic, antihypertensive and lipid-lowering agents), abundance of microbial species and pathways – using linear regression. We then used the residuals for QTL mapping. We considered a genome-wide P value $< 5 \times 10^{-8}$ to be the threshold for significant BA QTLs (baQTLs). We report all independent baQTLs (clumping variants with linkage disequilibrium $r^2 < 0.05$ and a 500kb window; Watanabe et al., 2017) with a P value $< 5 \times 10^{-8}$. Finally, we also evaluated genetic variants located in or near (within 250kb) confirmed BA metabolism-related genes at P value $< 1 \times 10^{-5}$ (63 genes in total, list available via: <https://github.com/GRONINGEN-MICROBIOME-CENTRE>).

Microbiome-wide correlation

To identify microbial BA determinants, we assessed correlations between the abundance levels of BA entities and microbial species and pathways. Microbial species and pathway datasets were also inverse-rank-transformed and corrected for sequencing depth and other covariates (age, sex, BMI, diseases and medications), as we had done with BA entities. To explore the correlations between the 39 BA entities and microbiome features (113 species and 335 pathways), we performed Spearman correlation tests and calculated false discovery rate (FDR) based on 1000 times permutation.

Independent effects of genetics and microbiome

Since genetics may impact the gut microbiome (Chen et al., 2018), we checked for this impact on the BA entities that associated with both genetics ($p < 5 \times 10^{-8}$) and the microbiome at FDR < 0.05 (after adjusting for the covariates of age, sex, BMI, disease and medication). In total, 16 BA entities associated to both genetics and microbiome. We checked the associations between independent SNPs and microbial traits that correlated with the same BA entity using Spearman correlation.

Correlation between disease and medication phenotypes

To check whether BA entities can be related to liver fat content, the presence of disease (diabetes, hypertension and dyslipidemia) and use of associated medications (antidiabetic, antihypertensive and lipid-lowering agents), we inverse-rank-transformed BA entities and adjusted them for confounding factors age and sex. Spearman correlation was then used to assess correlation strengths and 1000 times permutations were applied to calculate FDR.

Correlation with disease-relevant plasma metabolites

We inverse-rank-transformed plasma metabolites and corrected them for the age and sex. Spearman correlation was carried out to test the correlations between BA entities and disease-relevant plasma metabolites, and we calculated FDR based on 1000 times permutations.